Attorney Docket No.: A1992007DC2
Express Mailing Label No.: ET852159973US

A Patent Application

for

ELECTRONIC FILM EDITING SYSTEM USING BOTH FILM AND VIDEOTAPE FORMAT

By

Eric C. Peters
80 Carleton Road
Carlisle, Massachusetts 01741, USA
Citizen of the United States

Patrick D. O'Connor
84 Simpson Drive
Framingham, Massachusetts 01701, USA
Citizen of the United States

and

Michael E. Phillips
21 Bartlett Crescent
Brookline, Massachusetts 02146, USA
Citizen of the United States

ELECTRONIC FILM EDITING SYSTEM USING BOTH FILM AND VIDEOTAPE FORMAT

Background of the Invention

This invention relates to techniques for electronically editing film.

Film video and audio source material is frequently edited digitally using a computer system, such as the Avid/1 Media Composer from Avid Technology, Inc., of Tewksbury, Massachusetts, which generates a digital representation of a source film, allowing a film editor to edit the digital version, rather than the film source itself. This editing technique provides great precision and flexibility in the editing process, and is thus gaining popularity over the old style of film editing using a flatbed editor.

The Avid/1 Media Composer accepts a videotape version of a source film, created by transferring the film to videotape using the so-called telecine process, and digitizes the videotape version for editing via manipulation by computer. The operation of the Media Composer is described more fully in copending application U.S.S.N. 07/866,829, filed April 10, 1992, and entitled Improved Media Composer. The teachings of that application are incorporated herein by reference. Editing of the digitized film version is performed on the Media Composer computer using CRT monitors for displaying the digitized videotape, with the edit details being based on videotape timecode specifications. Once editing is complete, the Media Composer creates an edited videotape and a corresponding edit decision list (EDL) which documents the videotape timecode specification details of the edited videotape. The film editor uses this EDL to specify a cut and assemble list for editing the source film. While providing many advantages over the old style flatbed film editing technique, this electronic editing technique is found to be

cumbersome for some film editors who are unaccustomed to videotape timecode specifications.

Summary of the Invention

In general, in one aspect, the invention provides a system for generating a digital representation of a video signal comprised of a sequence of video frames which each include two video fields of a duration such that the video plays at a first prespecified rate of frames per second. The sequence of video frames includes a prespecified number of redundant video fields. In the invention, redundant video fields in the video frame sequence are identified by a video processor, and the video frame sequence is digitized by an analog to digital convertor, excluding the identified redundant video fields. The digitized video frames are then compressed by a video compressor to generate a digital representation of the video signal which plays at a second prespecified rate of frames per second.

In preferred embodiments, the invention further provides for storing the digitized representation of the video signal on a digital storage apparatus. The redundant video fields are identified by assigning a capture mask value to each video field in the video frame sequence, the capture mask value of a field being a "0" if the field is redundant, and the capture mask value of a field being a "1" for all other video fields. A video frame grabber processes the video frame sequence based on the capture mask values to exclude the identified redundant video frames from being digitized. The video compressor compresses the video frames based on JPEG video compression.

In other preferred embodiments, the first prespecified video play rate is 29.97 frames per second and the second prespecified digital video play rate is 24 frames per second. The rate of the analog video signal is increased from 29.97 frames per second to 30 frames per second before

the step of digitizing the video frame sequence. In further preferred embodiments, the analog video signal is a video representation of film shot at 24 frames per second, and the digital video play rate of 24 frames per second corresponds to the 24 frames per second film shooting rate. The analog video signal is a representation of film that is transferred to the video representation using a telecine apparatus.

In general, in another aspect, the invention provides an electronic editing system for digitally editing film shot at a first prespecified rate and converted to an analog video representation at a second prespecified rate. The editing system includes analog to digital converting circuitry for accepting the analog video representation of the film, adjusting the rate of the analog video such that the rate corresponds to the first prespecified rate at which the film was shot, and digitizing the adjusted analog video to generate a corresponding digital representation of the film. Further included is a digital storage apparatus for storing the digital representation of the film, and computing apparatus for processing the stored digital representation of the film to electronically edit the film and correspondingly edit the stored digital representation of the film.

In preferred embodiments, the system further includes digital to analog converting circuitry for converting the edited digital representation of the film to an analog video representation of the film, adjusting the rate of the analog video from the first prespecified rate to the second prespecified video rate, and outputting the adjusted analog video. Preferably, the analog video representation of the film accepted by the analog to digital converting circuity is an NTSC videotape. The apparatus for storing the digital representation of the film also stores a digitized version of a film transfer log corresponding to the digital representation of the film. The system includes display apparatus for

displaying the digitized version of the film as the film is electronically edited and displaying a metric for tracking the location of a segment of the film as the segment is displayed, the metric being based on either film footage code or video time code, as specified by the system user.

The electronic editing system of the invention allows users to provide the system with film formatted on standard videotapes, NTSC tapes, for example, and yet allows the video to be digitally edited as if it were film, i.e., running at film speed, as is preferred by most film editors. By reformatting the analog video as it is digitized, the system provides the ability to electronically edit film based on the same metric used in conventional film editing.

Brief Description of the Drawings

Fig. 1 is a schematic diagram of the electronic editing system of the invention.

Fig. 2 is a diagram of the telecine film-tape transfer pulldown scheme.

Fig. 3 is a schematic diagram of the telecine film-tape transfer system.

Fig. 4 is an Evertz Film Transfer Log produced by the telecine transfer system and processed by the editing system of the invention.

Fig. 5 is an illustration of a video screen showing the electronic bin generated by the editing system of the invention.

Fig. 6 is a diagram of the scheme employed by the editing system in digitizing a video input to the system.

Fig. 7 is an illustration of a video screen showing the digitized video to be edited on the electronic editing system of the invention.

Description of the Preferred Embodiment

Referring to Fig. 1, there is shown the electronic editing system of the invention 10, including two CRT displays 12, 14 for displaying

digitized film during an editing session, and an audio output device 16, for example, a pair of speakers, for playing digitized audio during an editing session. The displays 12, 14 and audio output 16 are all controlled by a computer 18. Preferably, the computer is a Macintosh™ II_{ci}, II_{fx}, Quadra 900, or Quadra 950 all of which are available from Apple Computer, Inc., of Cupertino, California. The system includes a video tape recorder (VTR) 20 for accepting an electronic version of film footage, which is preprocessed and digitized by a video analog to digital converter (A/D) 26. A timing circuit 28 controls the speed of the video being digitized, as described below. A video compressor 30 is connected to the video A/D for compressing the electronic image data to be manipulated by the computer 18. An audio A/D 22 and audio processor 24 process audio information from the electronic version of film footage in parallel with the video processing. Disc storage 32 communicates with the computer to provide memory storage for digitized electronic image data. This disc storage may be optical, magnetic, or some other suitable media. The editing system is user-interfaced via a keyboard 34, or some other suitable user control interface.

In operation, video and audio source material from a film which has been transferred to a videotape is received by the system via the video tape recorder 20, and is preprocessed and digitized by the audio A/D 22, audio processor 24, video A/D 26, and video compressor 30, before being stored in the disc storage 32. The computer is programmed to display the digitized source video on a first of the CRTs 12 and play the accompanying digitized source audio on the audio output 16. Typically source material is displayed in one window 36 of the first CRT 12 and edited material is displayed in a second window 38 of that CRT. Control functions, edit update information, and commands input from the keyboard 32 are typically displayed on the second system CRT 14.

Once a film is input to the system, a film editor may electronically edit the film using the keyboard to make edit decision commands. As will be explained in detail below, the electronic editing system provides the film editor with great flexibility, in that the video displayed on the system CRT 12 may be measured and controlled in either the domain of film footage or the domain of videotape time code. This flexibility provides many advantages over prior electronic editing systems. At the end of an editing session, the electronic editing system provides the film editor with an edited videotape and both tape and film edit command lists for effecting the edits from the session on film or videotape.

As explained above, the electronic editing system 10 requires a videotape version of a film for electronic manipulation of that film. Such a tape is preferably generated by a standard film-tape transfer process, the telecine process, which preferably uses the Time Logic Controller™ telecine (TLC), a device that converts film into a video signal, then records the signal on videotape. A TLC controls the film-tape transfer more precisely than non-TLC systems. In addition, it outputs a report, described below, that includes video format specifications, i.e., timecode, edge number, audio timecode, scene, and take for each reference frame in each tape, thereby eliminating the need to search through the video or film footage manually to find the data required for creating a log of video playing particulars. Other telecine systems may be used, however, depending on particular applications.

Transfer from film to tape is complicated by the fact that film and video play at different rates--film plays at 24 frames per second (fps), whereas PAL video plays at 25 fps and NTSC (National Television Standards Committee) video plays at 29.97 fps. If the film is shot at the standard rate of 24 fps and then transferred to 29.97 fps NTSC video, the difference between the film and video play rates is large (and

typically unacceptable). As a result, the film speed must be adjusted to accommodate the fractional tape speed, and some film frames must be duplicated during the transfer so that both versions have the same duration. However, if the film is shot at 29.97 fps, then transferring the footage to NTSC video is simple. Each film frame is then transferred directly to a video frame, as there are the same number of film and video frames per second.

Considering the most common case, in which 24 fps film is to be transferred to 29.97 fps NTSC videotape, the telecine process must provide both a scheme for slowing the film and a frame duplication scheme. The film is slowed down by the telecine apparatus by 0.1% of the normal film speed, to 23.976 fps, so that when the transfer is made, the tape runs at 29.97 fps, rather than 30 fps. To illustrate the frame duplication scheme, in the simplest case, and disregarding the film slow-down requirement, one second of film would be transferred to one second of video. The one second of film would include 24 frames of film footage, but the corresponding one second of video would require 30 frames of footage. To accommodate this discrepancy, the telecine process duplicates one film frame out of every four as the film is transferred to tape, so that for each second of film footage, the corresponding second of tape includes six extra frames.

Each video frame generated by the telecine process is actually a composite of two video fields: an odd field, which is a scan of the odd lines on a video screen, and an even field, which is a scan of the even lines. A video field consists of 262 1/2 scan lines, or passes of an electron beam across a video screen. To create a full video frame comprised of 525 scan lines, an odd field, or scan of the odd lines, is followed by an even field, or scan of the even lines. Thus, when a duplicate video frame is generated and added in the telecine process, duplicate video fields are

actually created. During play of the resulting tape, each two video fields are interlaced to make a single frame by scanning of the odd lines (field one) followed by scanning of the even lines (field two) to create a complete frame of NTSC video.

There are two possible systems for creating duplicate video fields in the telecine process, those systems being known as 2-3 pulldown and 3-2 pulldown. The result of the 2-3 pulldown process is schematically illustrated in Fig. 2. In a film-tape transfer using 2-3 pulldown, the first film frame (A in Fig. 2) is transferred to 2 video fields AA of the first video frame; the next film frame B is transferred to 3 video fields BBB, or one and one half video frames, film frame C is transferred to two video fields CC, and so on. This 2-3 pulldown sequence is also referred to as a SMPTE-A transfer. In a 3-2 pulldown transfer process, this sequence of duplication is reversed; the first film frame A would be mapped to 3 video fields, the next film frame B would be mapped to 2 video fields, and so on. This 3-2 pulldown sequence is also referred to as a SMPTE-B transfer. In either case, 4 frames of film are converted into 10 video fields, or 5 frames of video footage. When a 2-3 pulldown sequence is used, an A, B, C, D sequence in the original film footage creates an AA, BB, BC, CD, DD sequence of fields in the video footage, as shown in Fig. 2. The telecine process slows down the film before the frame transfer and duplication process, so that the generated video frames run at 29.97 fps.

Referring to Fig. 3, as discussed above, the telecine 36 produces a video signal from the film; the video is generated to run at 29.97 fps and includes redundant film frames from the pulldown scheme. NAGRA™ audio timecode is the typical and preferable system used with films for tracking the film to its corresponding audiotape. During the telecine process, a corresponding audio track 38 is generated based on the

NAGRA™ and is slowed down by 0.1% so that it is synchronized to the slowed film speed. The sound from the film audiotrack is provided at 60 Hz; a timing reference 40 at 59.94 Hz slows the audio down as required. Thus, the telecine process provides, for recordation on a videotape 48 via a videotape recorder 20, a video signal (V in the figure), corresponding audio tracks, A_1-A_n, and the audio timecode (audio TC).

A further film-tape correspondence is generated by the telecine process. This is required because, in addition to the difference between film and video play rates, the two media employ different systems for measuring and locating footage. Film is measured in feet and frames. Specific footage is located using edge numbers, also called edge code or latent edge numbers, which are burned into the film. For example, Kodak film provides Keykode™ on the film to track footage. The numbers appear once every 16 frames, or once every foot, on 35 mm film. The numbers appear once every 20 frames, or every half foot, on 16 mm film. Note that 35 mm film has 16 frames per foot, while 16 mm film has 40 frames per foot. Each edge number includes a code for the film manufacturer and the film type, the reel, and a footage counter. Frames between marked edge numbers are identified using edge code numbers and frame offsets. The frame offset represents the frame's distance from the preceding edge number.

Videotape footage is tracked and measured using a time-base system. Time code is applied to the videotape and is read by a time code reader. The time code itself is represented using an 8-digit format: XX:XX:XX:XX–hours:minutes:seconds:frames. For example, a frame occurring at 11 minutes, 27 seconds, and 19 frames into the tape would be represented as 00:11:27:19.

It is preferable that during the telecine conversion, a log, called a Film Transfer Log (FTL), is created that makes a correspondence

between the film length-base and the video time-base. The FTL documents the relationship between one videotape and the raw film footage used to create that tape, using so-called sync points. A sync point is a distinctive frame located at the beginning of a section of film, say, a clip, or scene, which has been transferred to a tape. The following information documents a sync point: edge number of the sync point in the film footage, time code of the same frame in the video footage, the type of pulldown sequence used in the transfer, i.e., 2-3 pulldown or 3-2 pulldown, and the pulldown mode of the video frame, i.e., which of the A, B, C, and D frames in each film five-frame series corresponds to the sync point frame.

As shown in Fig. 3, an Evertz 4015 processor accepts the video signal from the telecine and the audio TC corresponding to the audiotrack and produces a timecode based on a synchronization of the audio and video. Then an Evertz PC 44 produces an Evertz FTL 46 which includes the sync point information defined above.

Fig. 4 illustrates a typical Evertz FTL 46. Each column of the log, specified with a unique Record #, corresponds to one clip, or scene on the video. Of particular importance in this log is the VideoTape Time Code In (VTTC IN) column 50 and VideoTape Time Code Out (VTTC OUT) column 52. For each scene, these columns note the video time code of the scene start and finish. In a corresponding manner, the Keyin column 54 and Keyout column 56 note the same points in film footage and frames. The Pullin column 58 and Pullout column 60 specify which of the A, B, C, or D frames in the pulldown sequence correspond to the frame at the start of the scene and the close of the scene. Thus, the FTL gives scene sync information that corresponds to both the video domain and the film domain.

The electronic editing system of the invention accepts a videotape

produced by the telecine process and an Evertz FTL, stored on, for example, a floppy disk. When the FTL data on the disk is entered into the system, the system creates a corresponding bin in memory, stored on the system disc, in analogy to a film bin, in which film clips are stored for editing. The electronic bin contains all fields necessary for film editing, all comments, and all descriptions. The particulars of the bin are displayed for the user on one of the system's CRTs. Fig. 5 illustrates the display of the bin. It corresponds directly to the Evertz FTL. The "Start" and "End" columns of the bin correspond to the VideoTape Time Code In and VideoTape Time Code Out columns of the FTL. The "KN Start" and "KN End" columns of the bin correspond to the Keyin and Keyout columns of the FTL. During an editing session, the bin keeps track of the editing changes in both the video time-base and the film footage-base, as described below. Thus, the bin provides the film editor with the flexibility of keeping track of edits in either of the metrics.

Referring again to Fig. 1, when the electronic editing system 10 is provided with a videotape at the start of a film editing session, the videotape recorder 20 provides to the computer 18 the video and audio signals corresponding to the bin. The video signal is first processed by a video A/D coprocessor 26, such as the NuVista board made by TrueVision of Indianapolis, Indiana. A suitable video coprocessor includes a video frame grabber which converts analog video information into digital information. The video coprocessor has a memory which is configured using a coprocessor such as the TI34010 made by Texas Instruments, to provide an output data path to feed to the video compression circuitry, such as JPEG circuity, available as chip CL550B from C-Cube of Milpitas, California. Such a configuration can be performed using techniques known in the art. A timing circuit 28 controls the speed of the video signal as it is processed.

In operation, the video A/D 26 processes the video signal to reformat the signal so that the video represented by the signal corresponds to film speed, rather than videotape speed. The reformatted signal is then digitized, compressed, and stored in the computer for electronic film editing. This reformatting process allows users to provide the editing system with standard videotapes, in NTSC format, yet allows the video to be edited as if it were film, i.e., running at film speed, as is preferred by most film editors.

Referring also to Fig. 6, in this reformatting process, the speed of the video from the videotape is increased from 29.97 fps to 30 fps, as commanded by the timing circuitry 28 (Fig. 1). Then the fields of the video are scanned by the system, and based on the pulldown sequence and pulldown mode specified for each scene by the bin, the redundant video fields added by the telecine process are noted, and then ignored, while the other, nonredundant, fields are digitized and compressed into digital frames. More specifically, a so-called "capture mask" is created for the sequence of video fields; those fields which are redundant are assigned a capture value of "0" while all other fields are assigned a capture value of "1". The system coprocessor reads the entire capture mask and only captures those analog video fields corresponding to a "1" capture value, ignoring all other fields. In this way, the original film frame sequence is reconstructed from the video frame sequence. Once all the nonredundant fields are captured, the fields are batch digitized and compressed to produce digitized frames.

Assuming the use of the 2-3 pulldown scheme, as discussed above, in the capture process, the first two analog video fields (AA in Fig. 6) would each be assigned a capture value of "1", and thus would be designated as the first digital frame; the next two analog video fields BB would also each be assigned a capture value of "1", and be designated as

the second digital frame; but the fifth analog video field B, which is redundant, would be assigned a capture value of "0", and would be ignored, and so on. Thus, this process removes the redundant 6 frames added by the telecine process for each film second from the video, thereby producing a digitized representation which corresponds directly to the 24 fps film from which the video was made. This process is possible for either the 2-3 or 3-2 pulldown scheme because the bin specifies the information necessary to distinguish between the two schemes, and the starting frame (i.e., A, B, C, or D) of either sequence is given for each scene.

Appendix A of this application consists of an example of assembly language code for the MacIntosh™ computer and the TI 34010 coprocessor for performing the reformatting process. This code is copyrighted, and all copyrights are reserved.

Referring again to Fig. 1, an audio A/D 22 accepts audio from a videotape input to the editing system, and like the video A/D 26, increases the audio speed back to 100%, based on the command of the timing circuitry 28. The audio is digitized and then processed by the audio processor 24, to provide digitized audio corresponding to the reformatted and digitized video. At the completion of this digitization process, the editing system has a complete digital representation of the source film in film format, i.e., 24 fps, and has created a bin with both film footage and video timecode information corresponding to the digital representation, so that electronic editing in either time-base or footage-base may begin.

There are traditionally three different types of film productions that shoot on film, each type having different requirements of the electronic editing system. The first film production type, commercials, typically involves shooting on 35 mm film, transferring the film to a

videotape version using the telecine process, editing the video based on the NTSC standard, and never editing the actual film footage, which is not again needed after the film is transferred to video. Thus, the electronic editing is here preferably based on video timecode specifications, not film footage specifications, and an NTSC video is preferably produced at the end of the edit process. The electronic commercial edit should also preferably provide an edit decision list (EDL) that refers back to the video; the edited version of this video is typically what is actually played as the final commercial.

The second production type, episodic film, involves shooting on either 35 or 16 mm film, and producing an NTSC videotape version and additionally, an (optional) edited film version for distribution in markets such as HDTV (High Definition Television) or foreign countries. To produce the edited film footage for the film version, the film is transferred to videotape using the telecine process, and electronic editing of the film is here preferably accomplished based on film footage, and should produce a cutlist, based on film footage specifications, from which the original film is cut and transferred to the NTSC format. To produce a video version, the videotape is then preferably edited based on video timecode specifications to produce an EDL for creating an edited video version.

The third film production type, feature film, typically involves shooting on 35 mm film, and produces a final film product; thus electronic editing is here preferably based on film footage specifications to produce a cutlist for creating a final film version.

The user interface of the electronic editing system is designed to accommodate film editors concerned with any of the three film production types given above. As shown in Fig. 7, the video display CRT 12 of the system, which includes the source video window 36 and edited

video window 38, displays metrics 37, 39 for tracking the position of digital frames in a scene sequence currently being played in the source window or the edit window. These metrics may be in either film footage format or video time code format, whichever is preferred by the user. Thus, those film editors who prefer film footage notation may edit in that domain, while those film editors who prefer video timecode notation may edit in that domain. In either case, the digitized frames correspond exactly with the 24 fps speed of the original source film, rather than the 29.97 fps speed of videotape, so that the electronic edits produced by the electronic editing correspond exactly with the film edits, as if the film were being edited on an old-style flat bed editor.

As an example of an editing session, one scene could be selected from the bin and played on the source window 36 of the system CRT display 12. A film editor could designate frame points to be moved or cut in either timecode or film footage format. Correspondingly, audio points could be designated to be moved or the audio level increased (or decreased). When it is desired to preview a video version of such edits, an NTSC video is created by the system based on the sync information in the electronic bin, from the system disc storage, to produce either a so-called rough cut video, or a final video version. In this process, the system generates an analog version of the digital video signal and restores the redundant video frames necessary for producing the NTSC video rate. The system also produces a corresponding analog audio track and decreases the audio speed so that the audio is synchronized with the video. In this way, the system essentially mimics the telecine process by slowing down the video and audio and producing a 29.97 fps videotape based on a 24 fps source.

Referring again to Fig. 1, in creating an NTSC video from a digitized film version, the video compressor 30 retrieves the digitized

video frames from the computer 18 and based on the electronic bin information, designates video fields. The video A/D 26 then creates an analog version of the video frames and processes the frames using a pulldown scheme like that illustrated in Fig. 2 to introduce redundant video frames. The video speed is then controlled by the timing circuit 28 to produce 29.97 fps video as required for an NTSC videotape. Correspondingly, the system audio process 24 and audio A/D 22 processes the digital audio signal based on the electronic bin to generate an analog version of the signal, and then slows the signal by 0.1% to synchronize the audio with the NTSC video. The final video and audio signals are sent to the videotape recorder 20, which records the signals on a videotape.

The electronic editing system may be programmed to produce an edit listing appropriate to the particular media on which the finalized version of the film source material is to appear. If the source film material is to be finalized as film, the system may be specified to produce a cut list. The cut list is a guide for conforming the film negative to the edited video copy of the film footage. It includes a pull list and an assemble list. The assemble list provides a list of cuts in the order in which they must be spliced together on the film. The pull list provides a reel-by-reel listing of each film cut. Each of these lists specifies the sync points for the cuts based on film footage and frame keycode, as if the film had been edited on a flatbed editor. If the source film material is to be finalized as video, the system may be specified to produce an edit decision list (EDL). The EDL specifies sync points in video time code, as opposed to film footage. The editing system generates the requested edit lists based on the electronic bin; as the film is electronically edited, the bin reflects those edits and thus is a revised listing of sync points corresponding to the edited film version. Because

the bin is programmed to specify sync points in both film footage and video timecode, the editing system has direct access to either format, and can thereby generate the requested EDL or assemble and pull lists. Appendix B consists of examples of an EDL, assemble lists, and pull lists, all produced by the electronic editing system. Thus, at the end of an electronic film edit, the editing system provides a film editor with an NTSC videotape of the film edits and a edit list for either film or videotape.

Other embodiments of the invention are within the scope of the claims. What is claimed is: